Remarks

Reconsideration of the application is respectfully requested in view of the foregoing amendments and following remarks. With entry of amendments included herein, claims 1-2 and 4-37 are pending in this application. Claims 1, 9, 23, 31, 36, and 37 are independent. No claims have been allowed. Claims 1, 9, 23-24, 31, and 36 have been amended and claims 3, 10 and 17 have been canceled.

Claims Rejections Under 35 USC § 103

Claims 1-2, 4-8, 17, 20, and 23-30 are rejected under 35 U.S.C. 103(a) as being unpatentable over USP 5,729,471 *Jain* et al. ("*Jain*") in view of USP 5,612,743 ("Lee"). Applicants respectfully submit that the claims in their present form are allowable over the applied art. To establish a *prima facie* case of obviousness, three basic criteria must be met. First, there must be some suggestion or motivation, either in the references themselves or in the knowledge generally available to one of ordinary skill in the art, to modify the reference or to combine reference teachings. Second, there must be a reasonable expectation of success. Finally, the prior art reference (or references when combined) must teach or suggest all the claim limitations. (MPEP § 2142.).

Independent Claim 23

Amended claim 23 recites as follows:

> A method of recovering a three-dimensional scene from a sequence of two-dimensional frames, comprising:
>
> (a) identifying at least a first base frame in a sequence of two-dimensional frames;
>
> (b) adding the at least first base frame to create a first segment of frames of the sequence;
>
> (c) selecting feature points in at least the first base frame in the first segment of frames in the sequence;

(d) analyzing a next frame in the sequence to identify the selected feature points in the next frame;

(e) determining a number of the selected feature points from the base frame that are also identified in the next frame; and

(f) if the number of the selected feature points from the base frame that are also identified in the next frame is greater than or equal to a threshold number, adding the next frame to the first segment of frames of the sequence.

The applied references, *Jain* and *Lee*, both individually and in combination, fail to teach or suggest many aspects of claim 23. For instance, combining *Jain's* teaching of capturing video shot sequences at a standard 30 frames per second with *Lee's* teaching of comparing frames to a reference frame for determining pixel differential values for various regions within the frames fails to teach or suggest "*determining a number of the selected feature points from the base frame that are also identified in the next frame; and (f) if the number of the selected feature points from the base frame that are also identified in the next frame is greater than or equal to a threshold number, adding the next frame to the first segment of frames of the sequence*."

The Action relies on *Jain* and *Lee*. First of all, as the Action agrees, "*Jain* does not specifically disclose determining whether a threshold number of feature points from base frame are identified in the second frame, . . . adding the second frame to the segment." *See*, Action at Pg. 22, Lns. 7-10. Instead, the Action relies on *Lee* at Col. 2, Ln. 65- Col. 3, Ln. 3, which states as follows:

> subtracting the pixel value provided from the reference frame from a pixel value of said each pixel of the current frame to thereby provide **a differential pixel value;** (d) comparing on a pixel-by-pixel basis the differential pixel value with a threshold value TH and selecting one or more regions, each of the selected regions consisting of the pixels having their respective differential pixel values larger than the threshold value TH; (e) shifting the pixels within the selected regions to positions indicated by their respective motion vectors to thereby provide shifted regions; (f) detecting edge points from the reference frame; (g) determining none or more processing regions from the shifted regions, wherein the processing regions are the shifted regions which overlap with a portion of the edge points; (h) generating a first grid on the reference frame and

generating a second grid for each of the processing regions, wherein the second grid is formed by a portion of grid points of the first grid and newly added grid points, each of the newly added grid points being positioned at the center of a pair of neighboring grid points of the first grid in a horizontal or a vertical direction; and (i) selecting, as the feature points, a multiplicity of pixels in the reference frame based on the first and the second grids and the edge points.

First of all, *Lee* teaches determining pixel value differences between "selected regions" of frames by "selecting one or more regions, each of the selected regions consisting of the pixels having their respective **differential pixel values larger** than the threshold value TH." *See*, *Lee* at Col. 2, Ln .65- Col. 3, Ln. 31. However, what is claimed in Applicants' claim 23 is determining similarities between frames by tracking selected features to a base frame by "*determining whether a number of the selected feature points from the base frame that are also identified in the next frame; and if the number of the selected feature points from the base frame that are also identified in the next frame is greater than or equal to a threshold number, adding the next frame to the first segment of frames of the sequence.*" This is different than "comparing on a pixel-by-pixel basis the differential pixel value with a threshold value TH."

Furthermore, after the pixel-by-pixel evaluation described therein, nothing in *Lee* teaches or suggests "*adding the next frame to the first segment of frames of the sequence*" as recited by the Applicants' claim 23. Let alone, based on meeting the condition of "*if the number of the selected feature points from the base frame that are identified in the next frame is greater than or equal to the threshold number.*" In fact, nothing in *Lee* teaches or suggests building frame segments by tracking feature points. As Applicants understand *Lee,* it teaches "comparing on a pixel-by-pixel basis the differential pixel value with a threshold value TH" to designate ("pixel-by-pixel") different areas of the frames being compared as having conversion values of 0 or 1, which are later used for other processing. *See, e.g., Lee*, at Col. 5, Lns. 28-37. For instance, the Action points to FIG. 3 of *Lee,* but as Applicants understand *Lee*, FIG. 3 of *Lee* is related to edge detection and region selection within a frame being evaluated, wherein regions within the frame are selected based on the threshold comparison to a reference frame

(*i.e.*, those that are given a conversion value of "1") and are subjected to motion estimation, but *Lee* does not teach that the frame being evaluated in comparison to a reference frame is later added to a segment of frames of the original sequence that also contains the reference frame, based on results of the comparison. *See, generally, Lee* at Col. 5, Ln. 37- Col. 6, Ln. 12.

Accordingly, nothing in *Lee* teaches or suggests building frame segments by "*adding the next frame to the first segment of frames of the sequence*" the next frame being one frame in "*a sequence of two-dimensional frames*" which is compared to a "*base frame*" to determine whether the "*next frame*" should also be added to the segment comprising the "*base frame.*" Let alone, adding frames to frame segments based on whether the "*the number of the selected feature points from the base frame that are identified in the next frame is greater than or equal to the threshold number.*"

Since the applied references, do not teach or suggest at least one element of claim 23, claim 23 in its present form should be allowed.

Dependent Claim 24

Claim 24 recites as follows:

> The method of claim 23 further including if the number of the selected feature points from the base frame that are also identified in the next frame is less than the threshold number, adding the next frame to a second segment of frames of the sequence and designating the next frame that falls below the threshold number as a second base frame in a second segment.

Claim 24 depends on claim 23 and, thus, at least for the reasons set forth above with respect to claim 23, claim 24 should be allowed. However, claim 24 also recites independently patentable elements. For instance, *Jain* and *Lee*, both individually and in combination, fail to teach or suggest "*if the number of the selected feature points from the base frame that are also identified in the next frame is less than the threshold number, adding the next frame to a second segment of frames of the sequence and designating the next frame that falls below the threshold number as a second base frame in a second segment*" as recited in Applicants' claim

24. The Action at page 2 relies on *Lee* and states that "*Lee* also teaches that if a threshold number of feature points are identified in the second frame, adding the second frame to the segment. In figure 3, *Lee* suggests the cyclical process of determination of the threshold number values." *See,* Action at Pg. 2, Ln. 17-20. However, what is claimed in the amended claim 24 is entirely different. First of all, *Lee* fails to teach or suggest, "*designating the next frame that falls below the threshold number as a second base frame in a second segment.*" *Lee*'s comparison to a threshold value, described with respect to *Lee*'s FIG. 3 (and relied on by the Action) is stated in additional detail below at *Lee*, Col. 5, Lns. 18-37:

> The prediction signal is subtracted from the current frame signal at the subtractor 312, and the resultant data, i.e., a difference signal denoting the differential pixel values between the current frame signal and the prediction signal, is dispatched to a comparison block 313. The comparison block 313 compares on a pixel-by-pixel basis each of the differential pixel values included in the difference signal with a threshold value TH. The threshold value TH may be predetermined or determined adaptively according to the buffer occupancy, i.e., the amount of data stored in the buffer 109 shown in FIG. 1. If a differential pixel value is less than the threshold value TH, it is set to the conversion value 0. Otherwise, the differential pixel value is set to the conversion value 1. The conversion values are provided to a third frame memory 314. In FIG. 4, an error frame 41 formed by the conversion values stored in the third frame memory 314 is exemplarily shown. There are two distinct zones in the error frame 41: one is the regions (*e.g.,* A, B and C) with the conversion value 1; and the other, with the conversion value 0.

According to the paragraph above, all that *Lee* teaches is that if the "differential pixel value is less than the threshold value" the differential pixel value "is set to conversion value 0." Thus, nothing in *Lee* teaches or suggests that the "current frame" being evaluated with respect to "a reference frame" is added to "*a second segment of frames of the sequence*" and, more particularly, nothing in *Lee* teaches or suggests "*designating the next frame that falls below the threshold number as a second base frame in a second segment.*" In fact, FIG. 3 of *Lee* (on which the Action relies) shows the "reference frame signal" being fed into the block 310 independently of "current frame signal." Thus, nothing in *Lee* teaches or suggests that after

comparison at block 313, a "current frame signal" is ever designated as a "reference frame" for the next comparison let alone "*as a second base frame in a second segment*" as claimed in claim 24. However, the Applicants' specification at Pg. 10, Lns. 4-18 describes an exemplary description of the claimed segmenting process as follows:

> As indicated by arrow 102, each frame in an input sequence is taken in order and analyzed to determine whether it contains a threshold number of feature points tracked from the base frame or frames. If so, the frame is added to the current segment. At some point, however, a frame will not contain the threshold number of feature points and the decision made in decision 98 will be negative. At that point, a decision is made whether this is the final segment in the input sequence of images (decision 104). If it is the final segment (segment N), then the segmenting is complete (box 106). If however, there are more images in the input sequence, then the current segment is ended and the next segment is started (box 108). Typically, the last frame in the previous segment is also used as a base frame in the next segment. It also may be desirable to overlap several frames in each segment. Once the base frame is identified for the next segment, arrow 110 indicates that the process starts over for this next segment. Again, feature points are tracked with respect to the new base frame in the next segment. Because the number of frames depends on the feature points tracked between the frames, the segments can vary in length. (Emphasis added).

Furthermore, as noted above with respect to claim 23, *Lee* fails to teach any form of "*adding the next frame to a second segment of frames of the sequence.*" Let alone, based on "*if the number of the selected feature points from the base frame that are also identified in the next frame is less than the threshold number.*"

Since the applied references, do not teach or suggest at least one element of claim 24, claim 24 in its present form should be allowed.

Dependent claims 25-30

Claims 25-30 depend on claim 23 and, thus, at least for the reasons set forth above with respect to claim 23, claims 25-30 should be allowed.

Independent Claim 1

Amended claim recites as follows:

> A method of recovering a three-dimensional scene from two-dimensional images, the method comprising:
>
> providing a sequence of frames;
>
> dividing the sequence of frames into frame segments wherein the frames in the sequence comprise feature points and wherein the sequence of frames is divided into frame segments based upon frames in each frame segment having at least a minimum number of feature points being tracked to at least one base frame in the frame segment;
>
> performing three-dimensional reconstruction individually for each frame segment derived by dividing the sequence of frames; and
>
> combining the three-dimensional reconstructed segments together to recover a three-dimensional scene for the sequence of images.

The applied references, *Jain* and *Lee*, both individually and in combination, fail to teach or suggest many aspects of claim 1. For instance, combining *Jain's* teaching of capturing video shot sequences at a standard 30 frames per second with *Lee's* teaching of comparing "a current frame" to "a reference frame" for determining pixel differential values for various regions fails to teach or suggest *"dividing the sequence of frames into frame segments wherein the frames in the sequence comprise feature points and* ***wherein the sequence of frames is divided into frame segments based upon*** *frames in each frame segment having at least* ***a minimum number of feature points being tracked to at least one base frame in the frame segment."***

The Action relies on *Jain* and *Lee*. First of all, as the Action agrees, "*Jain* does not specifically disclose determining whether a threshold number of feature points from base frame are identified in the second frame, adding the second frame to the segment." *See*, Action at Pg. 22, Lns. 7-10. Instead, the Action relies on *Lee* comparing a "current frame" (pixel-by-pixel) to a "reference frame" to determine a "differential pixel value" and further processing including "the comparison block 313 compares on a pixel-by-pixel basis each of the differential pixel values included in the difference signal with a threshold value TH...if a

differential pixel value is less than the threshold value TH, it is set to the conversion value 0. Otherwise, the differential pixel value is set to the conversion value 1." *See, Lee* at Col. 5, Lns. 22-32. Although, *Lee* teaches comparing two frames, as a result of this comparison, "a differential pixel value (pixel-by-pixel)" is determined, which is not the same as ensuring "*frames in each frame segment having at least a minimum number of feature points being tracked to at least one base frame in the frame segment.*"

Moreover, nothing in *Lee* teaches or suggests that the result of its comparison operation is in any way to be used as criteria for *"dividing the sequence of frames into frame segments"* as claimed. *Lee* is simply silent as to "segments", since the frames in *Lee* are from a "video signal" stream and nothing in *Lee* teaches or suggests that this sequence is used for "*dividing the sequence of frames into frame segments.*" *Jain* on the other hand does teach segmenting a sequence of frames into an NTSC standard 30 frames per second. Thus, by combining the teachings of *Jain* with *Lee,* one would arrive at *Jain*'s fixed "standard" 30 frames per second frame segments, which is not based on any criteria and *Lee*'s method of comparing "a current frame and a reference frame." This fails to teach or suggest "*dividing the sequence of frames into frame segment"* **based on criteria** of *"each frame segment having at least a minimum number of feature points being tracked to at least one base frame in the frame segment."*

This is so at least because the frame segments of *Jain* are standard and not selected based on any criteria. *See, e.g.*, the Action at Pg. 16, Lns. 8-9, stating "every 30 frames obtained for each second, i.e., the standard NTSC frame rate (30 frames/sec), can be considered a segment..." Thus, by the Action's own logic, *Jain* applies no criteria at all for selecting a frame to belong to *Jain*'s 30 frame NTSC standard segment. Accordingly, one of ordinary skill in the art will fail to see a motivation for combining *Jain* which teaches segmenting that is not based on any criteria with *Lee,* which teaches a method of comparing two frames, but fails to teach or suggest that the result of the comparison, is a criteria for selecting which frames should be added to a segment by "*dividing the sequence of frames into*

frame segments" such that *"each frame segment having at least a minimum number of feature points being tracked to at least one base frame in the frame segment."*

Regardless, of the motivation for combining *Jain* with *Lee*, both references fail to teach or suggest all the claim limitations of claim 1. More particularly, the combination of *Jain* and *Lee* fails to teach or suggest *"dividing the sequence of frames into frame segment"* based on criteria that *"each frame segment having at least a minimum number of feature points being tracked to at least one base frame in the frame segment.* Instead, all that the combination of *Jain* and *Lee* teaches or suggests is *Jain*'s method of capturing video sequences at a standard NTSC 30 frame per second segment and *Lee*'s method of comparing a "current frame" to "a reference frame", which is not the same as applicant's claim 1 that recites *"dividing the sequence of frames into frame segment"* based on criteria that *"each frame segment having at least a minimum number of feature points being tracked to at least one base frame in the frame segment.*

Since the applied references, do not teach or suggest at least one element of claim 1, claim 1 in its present form should be allowed.

Dependent claims 2 and 4-8

Claims 2 and 4-8, ultimately depend on claim 1 and, thus, at least for the reasons set forth above with respect to claim 1, claims 2 and 4-8 should be also be allowed. Furthermore, each of the claims 2 and 4-8 also recite independently patentable features and, thus, should be allowed for that reason.

Dependent claim 17:

Claim 17 has now been canceled without prejudice and, thus, the rejection of claim 17 under 35 U.S.C. 103(a) is now moot.

Dependent claim 20:

Claim 20 depends on claim 9, which has been amended. Thus, at least for the reasons listed below with respect to claim 9, claim 20 recites at least one element that is not taught or suggested in either of the cited references, *Lee* and *Jain*. Thus, at least for this reason claim 20 should be allowed.

Independent claim 31:

Amended claim 31 recites as follows:

> In a method of recovering a three-dimensional scene from a sequence of two-dimensional frames, an improvement comprising dividing a long sequence of frames into segments and reducing the number of frames in each segment by representing the segments using between two and five representative frames per segment, wherein the representative frames are used to recover the three-dimensional scene and remaining frames are discarded so that the three-dimensional scene is effectively compressed, wherein dividing the long sequence into segments includes identifying a base frame and tracking feature points between frames in the sequence and the base frame and ending a segment whenever a frame does not contain a predetermined threshold of feature points that are contained in the base frame.

The applied references, *Jain* and *Lee*, both individually and in combination, fail to teach or suggest many aspects of claim 31. For instance, combining *Jain*'s teaching of capturing video shot sequences at a standard 30 frames per second with *Lee*'s teaching of comparing "a current frame" to "a reference frame" for determining pixel differential values for various regions fails to teach or suggest *"wherein dividing the long sequence into segments includes identifying a base frame and tracking feature points between frames in the sequence and the base frame and ending a segment whenever a frame does not contain a predetermined threshold of feature points that are contained in the base frame."* The Action relies on *Lee*. However, as noted above with respect to claims 1, 23, and 24, although *Lee* teaches comparing a "a current frame" to "a reference frame", all that *Lee* teaches is "comparing on a pixel-by-pixel basis the differential pixel value with a threshold value TH" to designate ("pixel-by-pixel") different areas of the frames being compared as having conversion values of 0 or 1."

See, e.g., Lee at Col. 5, Lns. 28- 37. More particularly, nothing in *Lee* teaches or suggests building a segment of frames, which comprises "*ending a segment whenever a frame does not contain a predetermined threshold of feature points that are contained in the base frame*" as recited in Applicant's claim 31, above.

Since the applied references do not teach or suggest at least one element of claim 31, claim 31 in its present form should be allowed.

Dependent claims 32-33 and 35:

Claims 32-33 and 35 depend on claim 31, which has been amended. Thus, at least for the reasons listed below with respect to claim 31, claims 32-33 and 35 recite at least one element that is not taught or suggested in either of the cited references, *Lee* and *Jain*. Thus, at least for this reason claims 32-33 and 35 should be allowed.

Claim Rejections 35 USC § 102

Claims 9-16, 18, 19, 21, 22, 36, and 37 are rejected under 35 U.S.C. 102(b) as being anticipated by *Jain*.

Independent claim 9:

Amended claim 9 recites as follows:

> A method of recovering a three-dimensional scene from two-dimensional images, the method comprising:
>
> identifying a sequence of two-dimensional frames that include two-dimensional images;
>
> dividing the sequence of frames into segments, wherein a segment includes a plurality of frames and wherein dividing includes, identifying a base frame, identifying feature points in the base frame; and determining the segments such that every frame in a segment has at least a predetermined percentage of feature points identified in the base frame;
>
> for each segment, encoding the frames in the segment into at least two virtual frames that include a three-dimensional structure for the segment and an uncertainty

associated with the segment and wherein encoding includes choosing at least two frames in the segment that are at least a threshold number of frames apart;

for each of the at least two chosen frames, projecting a plurality of three-dimensional points into a corresponding virtual frame; and

for each of the at least two chosen frames, projecting an uncertainty into the corresponding virtual frame.

The applied reference *Jain* fails to teach or suggest many aspects of Applicants' claim 9. For instance, *Jain* fails to teach or suggest *"determining the segments such that every frame in a segment has at least a predetermined percentage of feature points identified in the base frame."* The Action admits, "*Jain* does not specifically disclose determining whether a threshold number of feature points from base frame are identified in the second frame, adding the second frame to the segment." *See*, Action at Pg. 22, Lns. 7-10. Furthermore, for the reasons listed above with respect to claims 1 and 23, the other cited reference *Lee* both individually and in combination with *Jain* fails to teach or suggest *"determining the segments such that every frame in a segment has at least a predetermined percentage of feature points identified in the base frame."*

Again, both *Jain* and *Lee* individually and in combination fail to teach or suggest many other elements of claim 9. For instance, they fail to teach or suggest "dividing the sequence of frames into segments *...for each segment, encoding the frames in the segment into at least two virtual frames that include a three-dimensional structure for the segment and an uncertainty associated with the segment and wherein encoding includes choosing at least two frames in the segment that are at least a threshold number of frames apart; for each of the at least two chosen frames, projecting a plurality of three-dimensional points into a corresponding virtual frame; and for each of the at least two chosen frames, projecting an uncertainty into the corresponding virtual frame."* The Action relies on *Jain* teaching of 3D scene analysis to be performed on a set of 2D frames. *See*, *Jain* at Col. 22., Ln. 62 – Col. 23, Ln. 56. More particularly, *Jain* at Col. 23, Lns. 58 - Col. 24, line 4 states the following:

> Ideally the scene analysis process just described should be applied to every video frame in order to get the most precise information about (i) the location of players and (ii) the events in the scene. However, it would require significant human and computational effort to do so in the rudimentary, prototype, MPI video system because feature points are located manually, and not by automation. **Therefore, one key frame has been manually selected for every thirty frames, and scene analysis has been applied to the selected key frames**. For frames in between, player position and camera status is estimated by interpolation between key frames by proceeding under the assumption that coordinate values change linearly between a consecutive two key frames.

This fails to teach or suggest "*encoding the frames in the segment into at least two virtual frames...wherein encoding includes choosing at least two frames in the segment that are at least a threshold number of frames apart; for each of the at least two chosen frames, projecting a plurality of three-dimensional points into a corresponding virtual frame; and for each of the at least two chosen frames, projecting an uncertainty into the corresponding virtual frame*" as recited in amended claim 9. First of all, regarding *Jain,* the Action states "Jain discloses manually adjusting the number of key frames..." *See*, Action at Pg. 6, Lns. 16-18. Applicants disagree. What *Jain* teaches instead is "one key frame has been **manually selected for every thirty frames**, and scene analysis has been applied to the selected key frames." If according to the Action, *Jain*'s segment is the NTSC standard 30 frame segment, what *Jain* then teaches is that which one of 30 frames is to be subjected to 3D scene analysis is determined based on "**one key frame ...manually selected for every thirty frames.**" That is not the same as "*wherein encoding includes choosing at least two frames in the segment that are at least a threshold number of frames apart*" as recited in claim 9.

Additionally, nothing in *Jain* teaches or suggests that "*encoding the frames in the segment into at least two virtual frames*" is based on "choosing at least *two frames in the segment that are at least a threshold number of frames apart; for each of the at least two chosen frames, projecting a plurality of three-dimensional points into a corresponding virtual frame; and for each of the at least two chosen frames, projecting an uncertainty into the corresponding virtual frame.*" The Action seems to confuse the usage of the term "threshold

number" as used in the Applicants' claim 9 which relates to "a number of frames apart" with a "threshold" term of *Lee* which relates to something entirely different that is "differential pixel values."

Thus, the applied references, do not teach or suggest at least one element of claim 9, and as a result, claim 9 in its present form should be allowed.

Dependent claims 11-16 and 18-22:

Claims 11-16 and 18-22 depends on claim 9. Thus, at least for the reasons listed above with respect to claim 9, claims 11-16 and 18-22 recite at least one element that is not taught or suggested in either of the cited references, *Lee* and *Jain*. Thus, at least for this reason claims 11-16 and 18-22 should also be allowed.

Independent claim 36:

Amended claim 36 recites as follows:

> A computer-readable medium having computer-executable instructions for performing a method comprising:
>
> providing a sequence of two-dimensional frames;
>
> dividing the sequence into segments;
>
> calculating a partial model for each segment, wherein the partial model includes the same number of frames as the segment it represents and wherein the partial model includes three-dimensional coordinates and camera pose, the camera pose comprising rotation and translation, for features within the frames;
>
> extracting virtual key frames from each partial model, the virtual key frames having three-dimensional coordinates for the frames and an uncertainty associated with the frames; and
>
> bundle adjusting the virtual key frames to obtain a complete three-dimensional reconstruction of the two-dimensional frames.

Applied reference *Jain* does not teach or suggest many aspects of claim 36. For instance, *Jain* fails to teach or suggest "*calculating a partial model for each segment, wherein the partial model includes the same number of frames as the segment it represents and wherein*

the partial model includes three-dimensional coordinates and camera pose, the camera pose comprising rotation and translation, for features within the frames; extracting virtual key frames from each partial model, the virtual key frames having three-dimensional coordinates for the frames and an uncertainty associated with the frames; and bundle adjusting the virtual key frames to obtain a complete three-dimensional reconstruction of the two-dimensional frames."

First of all, *Jain* fails to teach or suggest "*calculating a partial model for each segment, wherein the partial model includes the same number of frames as the segment it represents."* The Action relies on FIG. 12 of *Jain* teaching "Image to ground projections." *See,* Action at Pg. 14, 2nd Para. *Jain* does not define or explain what these "Image to ground projections" are, but Applicants understand it to mean that a user selected "dynamic object" of interest is projected on the ground image to ascertain its relative location to a known point (*e.g.*, projecting a user selected player on the field markings on a football field). Thus, *Jain*'s "image to ground projections" have nothing to do with creating "partial models" let alone "*calculating a partial model for each segment, wherein the partial model includes the same number of frames as the segment it represents"* as claimed 9. Thus, for instance, nothing in *Jain* suggests "image to ground projections" have the "*includes the same number of frames as the segment it represents."*

Furthermore, for the sake of argument, if according to the Action, "image to ground projections" of *Jain*'s FIG. 12 is the same as Applicants' claimed "*partial model for each segment, wherein the partial model includes the same number of frames as the segment it represents"*, then *Jain* clearly fails to teach or suggest "*extracting virtual key frames from each partial model."* This so at least because, according to the Action, extraction of "key frames" in *Jain* is taught by *Jain*'s teaching that "one key frame has been manually selected for every thirty frames", which refers to the extraction of "one key frame...for every thirty frames" from the raw feed from the video camera (*e.g.*, *See*, *Jain* at FIG. 12). Nothing, in *Jain* teaches or suggests that the extraction of *Jain*'s "one key frame" is from the "image to ground

projection." Thus, it is evident that *Jain* completely fails to teach or suggest an entire step in claim 36, which recites, "*calculating a partial model for each segment, wherein the partial model includes the same number of frames as the segment it represents... extracting virtual key frames from each partial model.*"

Thus, at least for the reasons listed above, claim 36 recites at least one element that is not taught or suggested in either of the cited references, *Lee* or *Jain* or combination thereof for that matter. Thus, at least for this reason, claim 36 should be allowed.

Independent claim 37:

Claim 37 recites as follows:

> An apparatus for recovering a three-dimensional scene from a sequence of two-dimensional frames by segmenting the frames, comprising:
> means for capturing two-dimensional images;
> means for dividing the sequence into segments;
> means for calculating a partial model for each segment that includes three-dimensional coordinates and camera pose for features within the frames;
> means for extracting virtual key frames from each partial model; and
> means for bundle adjusting the virtual key frames to obtain a complete three-dimensional reconstruction of the two-dimensional frames.

The applied reference *Jain* fails to teach or suggest many aspects of Applicants' claim 37. For instance, *Jain* fails to teach or suggest *"means for calculating a partial model for each segment that includes three-dimensional coordinates and camera pose for features within the frames; means for extracting virtual key frames from each partial model."* As noted above with respect to claim 36, if according to the Action, "image to ground projections" of *Jain*'s FIG. 12 is the same as Applicants' claimed "*a partial model for each segment*", then *Jain* clearly fails to teach or suggest, "*extracting virtual key frames from each partial model.*" This is so at least because, according to the Action, extraction of "key frames" in *Jain* is taught by *Jain*'s teaching that "one key frame has been manually selected for every thirty frames" from

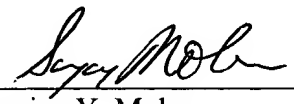
the raw feed from the video camera (*e.g.*, *See*, *Jain* at FIG. 12). Nothing, in *Jain* teaches or suggests that extraction of *Jain*'s "one key frame" is from the "image to ground projection."

Thus, at least for the reasons listed above, claim 37 recites at least one element that is not taught or suggested in either of the cited references, *Lee* or *Jain* or the combination thereof, for that matter. Thus, at least for this reason, claim 37 should be allowed.

Respectfully submitted,

KLARQUIST SPARKMAN, LLP

One World Trade Center, Suite 1600
121 S.W. Salmon Street
Portland, Oregon 97204
Telephone: (503) 595-5300
Facsimile: (503) 595-5301

By 
Sunjay Y. Mohan
Registration No. 56,739